



Paudyal, P., Shen, S., Yan, S., & Simeonidou, D. (2021). Toward Deployments of ML applications in Optical Networks. *IEEE Photonics Technology Letters*, 33(11), 537-540. [9410240].
<https://doi.org/10.1109/LPT.2021.3074586>

Peer reviewed version

Link to published version (if available):
[10.1109/LPT.2021.3074586](https://doi.org/10.1109/LPT.2021.3074586)

[Link to publication record in Explore Bristol Research](#)
PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via Institute of Electrical and Electronics Engineers at <https://ieeexplore.ieee.org/document/9410240> . Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

Toward Deployments of ML applications in Optical Networks

Paurakh Paudyal, Sen Shen, Shuangyi Yan*, and Dimitra Simeonidou, *Fellow, IEEE*

Abstract—To support the emerging 5G applications and the 5G bearer networks, optical networks, as the critical infrastructure, are continuously evolving to be more dynamic and automatic. The vision of future autonomous networks with low link margins requires precise estimation/prediction of the quality of transmission (QoT) of optical links. Machine learning (ML) technologies provide promising solutions to predict QoT of unestablished links. In this paper, we investigated hybrid modelling and transfer learning to address the key issues for deployment of ML applications in optical networks. The proposed approach for multiple-channel prediction reduces the training data requirement by 80% while obtaining the same MSE of 0.267dB compared with the model without transfer learning. The approach facilitates a streamlined ML life-cycle for data collection, training, and deployment.

Index Terms—Transfer learning, optical networks, Quality of Transmission estimation.

I. INTRODUCTION

The emerging 5G networks that were designed to support high-capacity network applications will bring an unprecedented amount of dynamic traffic to the underlying optical network infrastructures [1]. Future optical networks will need to be evolved to be more dynamic, with the ability to establish network connections with reduced margins to improve hardware utilisation [2]. Consequently, precise information about the quality of transmission (QoT) of the unestablished light paths as well as the impact of newly established light paths on the previous channels is of vital importance to operating low-margin optical networks efficiently.

With the unparalleled combination of high accuracy and low computational complexity in inference, Machine Learning (ML) based approaches have been explored to provide promising solutions in QoT estimation with either synthetic data [3] or pre-collected network operation data [4]. A quality of transmission (QoT) estimator based on support vector machines (SVM) reduced the essential computing time to evaluate the QoT of an established lightpath [5]. Multi-channel Q-factor prediction based on artificial neural networks (ANN) is investigated with real-time network operation and configuration information and the proposed ANN-based regression

model shows an accurate prediction of the Q-factor in both new unseen channels and existing channels [4]. The Reservoir Computing (RC)-based QoT prediction method can achieve an acceptable accuracy faster than ANN and SVM [6]. These solutions based on ML face big challenges in scalability as training and inferencing of ML models are carried out on the same network.

Recently, transfer learning (TL) [7] has been used to solve the scalability of the ANN-based QoT prediction [8], [9] in optical networks. Transfer learning allows retraining the pre-trained model with less data, which avoids comprehensive data collection in the deployed environment and can be combined with other ML algorithms, such as ANN and SVM. In [8], the proposed approach achieves Q-factor prediction accuracies with 0.42dB, 0.37dB and 0.67dB respectively, over three optical systems with a small number of training samples. CNN-based multi-impairment diagnosis technique with deep transfer learning has been proposed [10] and the presented deep TL-based method decreases the training time by more than 95% without sacrifice of the accuracy of 99.88%. In addition, deep neural networks (DNN) combined with transfer learning [11] is used to make quality of transmission (QoT) prediction in multi-domain optical networks to realize resource-efficient service supplying. Evaluation with experimental data demonstrates that the proposed approach can dramatically reduce the amount of required training data for new tasks with the same estimation accuracy. The above work proves that the combination of transfer learning and supervised learning is an effective method, which can save training time using only a few samples without sacrifice of the accuracy.

However, there are still several challenges that need to be addressed before their widespread deployment becomes realities. The large amounts of data that are required for ML model training are not yet available, particularly during the early phase of the fibre life-cycle when network monitoring data is lacking. Additionally, ML models must be responsive to system changes caused by component wear and ageing, and constantly evaluate their own efficacy so as to prevent inadequate quality of service. A coherent life-cycle for ML models is required to formalise the process of designing, testing, and deploying ML models in optical networks. To address this issue, reliable simulated datasets collected from the validated Gaussian noise model can be used on the initial model for pre-training and then the model can be retrained with far less practical data, which can achieve a extremely high precision.

In this letter, we propose a streamlined ML life-cycle for optical networks which utilises TL to combine synthetic data and practical data observed from a field-trial testbed. The work

This work was supported by Engineering and Physical Sciences Research Council (EPSRC) of UK through the TOUCAN project under Grant EP/L020009/1 and the NDFF project under Grant EP/S028854/1.

Paurakh Paudyal is with the Faculty of Engineering, University of Bristol, Bristol BS8 1UB, UK.

Sen Shen, Shuangyi Yan and Dimitra Simeonidou are with High Performance Networks Group (HPN), Smart Internet Lab, University of Bristol, Bristol BS8 1UB, UK. (e-mail:shuangyi.yan@bristol.ac.uk)

Copyright (c) 2019 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.

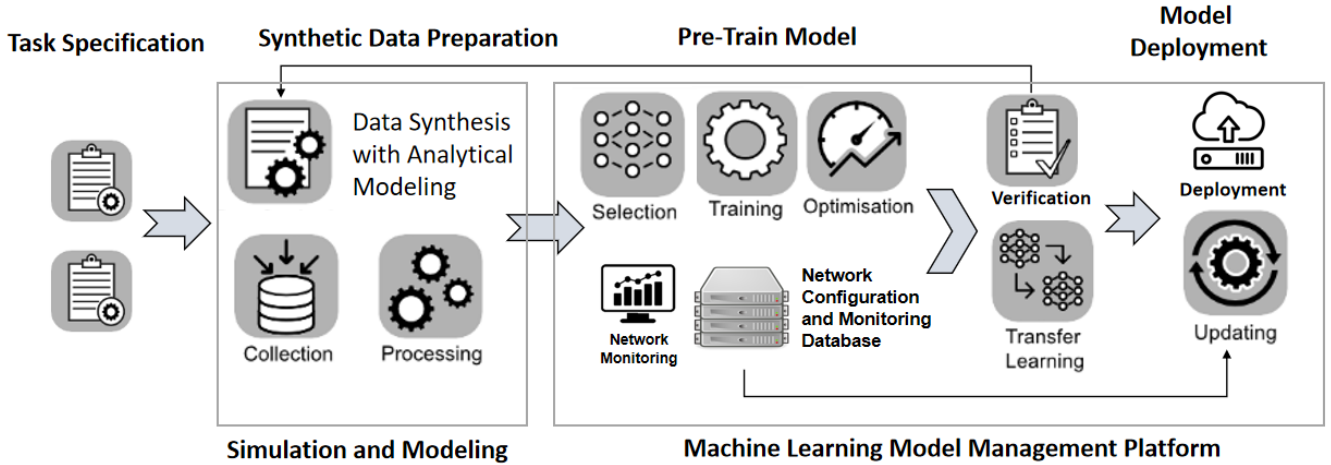


Fig. 1: Hybrid model life-cycle management for machine learning algorithms in optical networks

extended our work [12], by adding literature review of the management platform for machine learning model deployment and extra information about the experiments. Synthetic data gathered through coarse analytical modelling is used to obtain a QoT-prediction model with acceptable precision in the absence of practical network data. The QoT prediction model is then retrained and fine-tuned to achieve high precision prediction with practical data. The performance of the TL assisted ANN is evaluated by comparing it to a baseline ANN trained from scratch. The TL-assisted ANN achieves an MSE of 0.267 dB, equal to that of the baseline ANN, despite being trained on only 20% (200 samples) of the practical data used to train the baseline model. The training time is reduced for the TL assisted ANN, taking 6.67s in comparison to 19.47s for the baseline ANN. Our proposed approach reduces the volume of practical data required to train an ANN for QoT prediction, and facilitates rapid training and deployment of these predictors in future commercial optical networks. The streamlined ML lifecycle creates connections between synthetical data and practical data, and therefore, provides a possible approach for future deployment of machine learning models.

II. ML LIFE-CYCLE TOWARD DEPLOYMENTS IN OPTICAL NETWORKS

Our proposed four-phase approach for the ML life-cycle is outlined in Figure 1. Generally the process of ML has four parts, which are data collection, data transform, (re) training model and deployment. Transfer learning is used to connect available knowledge and practical network status. Firstly, a source learner is trained on synthetic data gathered offline through available modelling or simulation tools. The design phase consists of model selection, training, and optimisation. Initial values for hyper-parameters are set arbitrarily and optimised by performing a Grid Search across many hyper-parameter combinations. Then, the parameters of the trained source learner are transferred to the target (TL assisted) learner, which is fine-tuned with practical monitoring

data from the optical network. In our work, the time series database InfluxDB is used to store our data, which can have an organised management of time series data. This approach achieves convergence faster than training from scratch while reducing the required amount of training and validation data. The third phase is model validation, which aims to evaluate the model's ability to generalise for new, unseen network operations data. The results of this stage determine margins that must be utilised during deployment to ensure adequate quality of service. Successful models can be integrated into the software-defined networking (SDN) controller for autonomous and dynamic lightpath allocation. A remodelling algorithm accounting for all potential sources of system changes (e.g., component wear and ageing) can be implemented with practical network validation. With the proposed ML model management, network operation scenarios with complex topology and dynamic channel configurations can be offered ML models for QoT predictions.

The ML life cycle requires complex model management with a deep understanding of ML algorithms and models and expert knowledge in the application areas [13]. In this work, the ML model management platform *Weights & Biases* [14] is explored to manage the ML model as it has a clear view of the training results and shows the utilization information of CPU, system memory and detailed parameters about model structure. As for model deployment, the concept of MLaaS (Machine Learning as a service) intends to provide an automatic or semi-automatic cloud platform with integrated basic functions of Machine Learning implementation, such as data pre-processing, model training, model assessment and prediction. Amazon Machine Learning Services (AWS), Microsoft Azure Machine Learning, and Google Cloud AI are three of the leading Cloud MLaaS services that allow rapid model training and deployment with little or no data science expertise. We had a first attempt on Microsoft Azure ML, which only required to upload our corresponding datasets and choose the model we want to use. Microsoft Azure ML provides relatively complete tools for data pre-processing and other basic ML models,

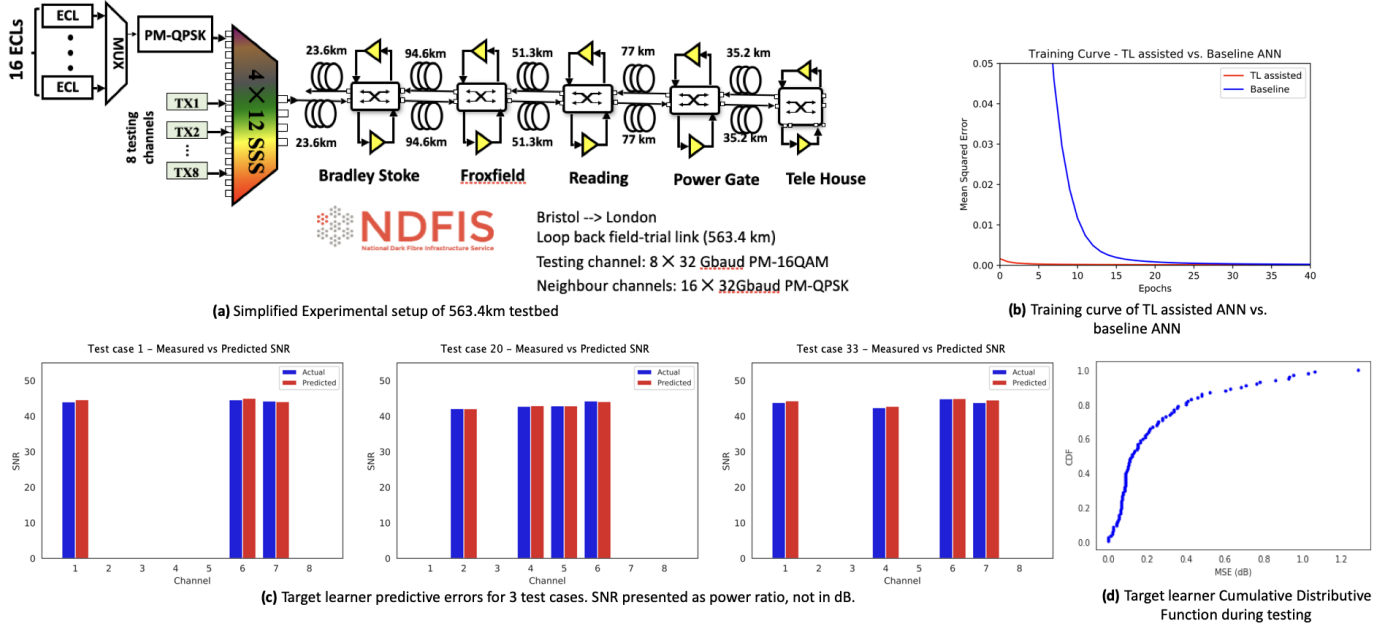


Fig. 2: Experimental setup and results

and we can upload our models written in R or Python to execute. However, extra knowledge about the application area is required for the practical deployment.

In the following section, we demonstrate the first three phases for ML model provision.

III. TRANSFER LEARNING BASED QOT ESTIMATOR

We utilise TL to reduce the required training data and increase speed of convergence for our ANN model. A source domain, \mathcal{D}_S , and source task, \mathcal{T}_S are defined. \mathcal{D}_S is comprised of the feature space \mathcal{X}_S and a marginal probability distribution $P_S(x)$. \mathcal{T}_S is comprised of the target space, \mathcal{Y}_S and the predictive function $f(\cdot)$. Similarly, a target domain, \mathcal{D}_T , and target task, \mathcal{T}_T are also defined. For our purposes, \mathcal{D}_S is the synthetic data set and \mathcal{D}_T is the practical network data set. TL aims to improve the learning of the target predictive function $f(\cdot)$ in \mathcal{D}_T using the knowledge in \mathcal{D}_S and \mathcal{T}_S , where $\mathcal{D}_S \neq \mathcal{D}_T$ but $\mathcal{T}_S = \mathcal{T}_T$ (both tasks are OSNR estimation). We utilise parameter transfer [9], and transfer a certain amount of parameters (weights, biases etc.) from the source learner to the target learner. This can be explained mathematically by:

$$w_S = w_0 + v_S \text{ and } w_T = w_0 + v_T \quad (1)$$

where w_S and w_T are parameters of the ANNs used for the source task and the target task respectively. w_0 is the set of parameters shared between both tasks, while v_S and v_T are task-specific parameters. The cost function for the target model can then be written:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - f(x_i; w_0, v_T))^2 \quad (2)$$

Where y_i is the actual output for the i^{th} data point and $f(x_i)$ is the output predicted by the target learner. Hence, parameter transfer induces faster training of the target model by making w_0 known. The two step process therefore is comprised of the

initial training of the source learner ANN on synthetic data, followed by the transferring of parameter knowledge to the target learner which is fine-tuned to practical network data.

Synthetic data was gathered on a simulated version of the national dark fibre facility (NDF) network using an open-source simulation software GNPpy [15], which is implemented based on the Gaussian noise (GN) analytical model [16]. GNPpy is an open source application for physical layer aware open optical networks defined by the Open Optical Packet Transport-Physical Simulation Environment (OOPt-PSE) group within the Telecom Infra Project (TIP) [17]. It provides a reliable software for the abstraction of data transport in optical networks and has been validated with a satisfactory accuracy in the experiments [18]. We use the optical network topology depicted in Figure 2a to provide GNPpy with the input and network elements. 16 external cavity lasers (ECLs) are combined together to generate 32 Gbaud PM-QPSK signals. 8 testing channels are generated by eight real-time PM-16QAM transponders. The NDF link from Bristol to London is comprised of five nodes with the loop-back configuration at the Telehouse node. Each node deploys two EDFAs with constant gain to compensate for the loss of the bidirectional link, and the generated optical signals are transmitted in a 563.4 km loop back field-trial link. The training data-set was comprised of 9000 synthetic data points generated by GNPpy, whilst the test data-set consisted of 1000 real data points gathered from a field-trial testbed.

Then, we built on our previous work in [4] and used the ANN architecture as a starting point for the source learner. The feature space was comprised of launch power, EDFA input/output power pairs for every node, and the channel coding vector, resulting in a 43 dimensional feature space. There are 8 outputs corresponding to the total OSNR for each of the channels at the end of the link. Channel coding is achieved through hot coding, where a vector of [1 1 0 0 1 0 0 0]

TABLE I: The summary of ANN model

Layer	Shape	Activation Function	Parameters
Input	43	-	-
Hidden 1	40	ReLU	1760
Hidden 2	20	ReLU	820
Output	8	Sigmoid	168

represents channel 1,2 and 5 being switched on. The ANN can thus infer the significance of a 0 in the channel vector, and predict a 0 for the OSNR of that channel. Similarly, the EDFA input/ output powers behave as a hot coded route vector by the inclusion of 0 values for all EDFAs that were not crossed on a given path, allowing the ANN to make inferences based on the route taken. We used a rectified linear unit (ReLU) function $f(x) = \max(0, x)$ as the activation for all neurons except those in the output layer, where the sigmoid function was used. An Adaptive Moment Estimation (Adam) was chosen as a stochastic gradient descent algorithm. The architecture of the ANN with two hidden layers shown in Table 1 was [43,40,20,8]. Batch size and number of epochs were 16 and 150 respectively. The hyper-parameters and architecture of the source learner are transferred to the target learner, which is re-trained on 200 random samples of practical data. After testing multiple parameter transfer schemes, fine-tuning (all parameters transferred) was chosen as the best approach, indicating a heavy correlation between \mathcal{T}_S and \mathcal{T}_T .

IV. RESULTS AND DISCUSSION

The performance of the target learner was evaluated against a baseline model trained from scratch with 500% more practical data (1000 samples). The training curves for the target learner and baseline model can be seen in Figure 2b. The target learner obtained convergence quicker than the baseline model, requiring only 2 epochs as opposed to 24. Training time was also reduced from 19.47s for the baseline model to 6.67s for the target on a personal computer (i7-10850H, 2.70GHz, 16GB RAM). Both models achieved an MSE of 0.27dB across 100 test data points and a 90th percentile accuracy of 0.704dB. The values of R2 score are 0.9963 for the target learner and 0.9957 for the baseline model respectively, which means that both of them have accurate estimations. The performance of the target learner for 3 test cases is shown in Figure 2c, and a plot of the cumulative distribution function (CDF) can be seen in Figure 2d. Our results indicate that robust QoT predictors can be trained with far less practical data than previously thought necessary. The current model is designed for a single link. The proposed ML lifecycle management supports the synthetic data generation, model pre-training, practical data collection, and model training. The work can be scaled up to the whole networks by preparing a multi-channel QoT prediction model for each link. A streamlined ML lifecycle management will be critical for training and deployment of ML based QoT predictors in dynamic optical networks.

V. CONCLUSION

In this paper, a streamlined ML life-cycle is presented toward future deployment in commercial optical networks. Based on the life-cycle, a TL-based QoT estimation is implemented with reduced training data requirements and faster

training times. Our approach paves the way for large-scale, rapid deployment of QoT predictors with complex network operation scenarios, to fully support dynamic optical networks during the 5G era.

REFERENCES

- [1] M. Fiorani, P. Monti, B. Skubic, J. Mårtensson, L. Valcarengi, P. Castoldi, and L. Wosinska, "Challenges for 5g transport networks," in *Proc. of the 8th IEEE Conference on Advanced Networks and Telecommunication Systems (ANTS)*, pp. 1–6, 2014.
- [2] Y. Pointurier, "Design of low-margin optical networks," *IEEE/OSA Journal of Optical Communications and Networking*, vol. 9, no. 1, pp. A9–A17, 2017.
- [3] C. Rottondi, L. Barletta, A. Giusti, and M. Tornatore, "Machine-learning method for quality of transmission prediction of unestablished light-paths," *IEEE/OSA Journal of Optical Communications and Networking*, vol. 10, no. 2, pp. A286–A297, 2018.
- [4] Z. Gao, S. Yan, J. Zhang, M. Mascarenhas, R. Nejabati, Y. Ji, and D. Simeonidou, "Ann-based multi-channel qot-prediction over a 563.4-km field-trial testbed," *Journal of Lightwave Technology*, vol. 38, no. 9, pp. 2646–2655, 2020.
- [5] I. Mata, I. de Miguel, R. J. Durán, J. C. Aguado, N. Merayo, L. Ruiz, P. Fernández, R. M. Lorenzo, and E. J. Abril, "A svm approach for lighthpath qot estimation in optical transport networks," in *2017 IEEE International Conference on Big Data (Big Data)*, pp. 4795–4797, 2017.
- [6] Y. Li, J. Wang, D. Cai, Z. Li, D. Fu, and L. Qin, "Towards high-efficient qot prediction of wide-area optical backbone network: A reservoir computing view," in *2020 IEEE/CIC International Conference on Communications in China (ICCC)*, pp. 384–388, 2020.
- [7] F. Zhuang, Z. Qi, K. Duan, D. Xi, Y. Zhu, H. Xiong, and Q. He, "A comprehensive survey on transfer learning," *Proceedings of the IEEE*, vol. 109, no. 1, pp. 43–76, 2021.
- [8] W. Mo, Y. Huang, S. Zhang, E. Ip, D. C. Kilper, Y. Aono, and T. Tajima, "Ann-based transfer learning for qot prediction in real-time mixed line-rate systems," in *2018 Optical Fiber Communications Conference and Exposition (OFC)*, pp. 1–3, 2018.
- [9] L. Xia, J. Zhang, S. Hu, M. Zhu, Y. Song, and K. Qiu, "Transfer learning assisted deep neural network for osnr estimation," *Optics express*, vol. 27, no. 14, pp. 19398–19406, 2019.
- [10] Y. Xu, D. Wang, M. Zhang, X. Zhou, Z. Zhang, J. Li, Y. Zhu, P. Xie, and N. Paerhati, "Deep transfer learning based multi-impairment diagnosis for pam-4 optical communication systems," in *2019 18th International Conference on Optical Communications and Networks (ICOON)*, pp. 1–3, 2019.
- [11] C. Liu, X. Chen, R. Proietti, and S. J. Ben Yoo, "Evol-tl: Evolutionary transfer learning for qot estimation in multi-domain networks," in *2020 Optical Fiber Communications Conference and Exhibition (OFC)*, pp. 1–3, 2020.
- [12] P. Paudyal, S. Yan, and D. Simeonidou, "Toward deployment of ml in optical networks, transfer learning, monitoring and modelling," in *Asia Communications and Photonics Conference/International Conference on Information Photonics and Optical Communications 2020 (ACP/IPOC)*, p. S4C.2, Optical Society of America, 2020.
- [13] S. Schelter, F. Bießmann, T. Januschowski, D. Salinas, S. Seufert, and G. Szarvas, "On challenges in machine learning model management," *IEEE Data Eng. Bull.*, vol. 41, pp. 5–15, 2018.
- [14] <https://wandb.ai/site>.
- [15] <https://gnpy.readthedocs.io/en/master/>.
- [16] P. Poggiolini, "The gn model of non-linear propagation in uncompensated coherent optical systems," *Journal of Lightwave Technology*, vol. 30, no. 24, pp. 3857–3879, 2012.
- [17] A. Ferrari, M. Filer, K. Balasubramanian, Y. Yin, E. Le Rouzic, J. Kundrat, G. Grammel, G. Galimberti, and V. Curri, "Gnpy: an open source application for physical layer aware open optical networks," *IEEE/OSA Journal of Optical Communications and Networking*, vol. 12, no. 6, pp. C31–C40, 2020.
- [18] A. Ferrari, J. Kundrat, E. L. Rouzic, M. Filer, A. Campanella, A. D'Amico, K. Balasubramanian, Y. Yin, O. Havliš, M. Hažlinský, J. Vojtěch, J. L. Augé, G. Grammel, G. Galimberti, and V. Curri, "The gnpy open source library of applications for software abstraction of wdm data transport in open optical networks," in *2020 6th IEEE Conference on Network Softwarization (NetSoft)*, pp. 386–394, 2020.